2015

# Wavelet-Coupled Machine Learning Methods for Drought Forecast Utilizing Hybrid Meteorological and Remotely-Sensed Data

R. Tan
*Portland State University*

Marek Perkowski
*Portland State University*, marek.perkowski@pdx.edu



Citation Details

R. Tan and M. Perkowski. (2015). Wavelet-Coupled Machine Learning Methods for Drought Forecast Utilizing Hybrid Meteorological and Remotely-Sensed Data. In Proc. Conference on Data Mining, DMIN15.

# Wavelet-Coupled Machine Learning Methods for Drought Forecast Utilizing Hybrid Meteorological and Remotely-Sensed Data

**R. Tan, and M. Perkowski**
Department of Electrical and Computer Engineering, Portland State University, Portland, Oregon, USA

**Abstract -** *In this study, a statistical drought early warning method is proposed using novel machine learning algorithms, with the inclusion of multiple drought-related attributes from precipitation, satellite-derived land cover vegetation indices, and surface discharge. The forecast is made for the long-term hydrological drought in the region of Central Valley, California. The wavelet transform analysis is employed in combination with support vector regression and artificial neural network algorithms for improving the drought prediction effectiveness. The performance of the drought prediction is evaluated using three statistical metrics: Coefficient of Determination ($R^2$), Root-Mean-Square Error (RMSE), and Mean-Absolute-Error (MAE). The results clearly indicate that using hybrid precipitation and satellite remotely-sensed data, the proposed wavelet-coupled machine learning method can effectively predict long-term drought in the area of Central Valley California, over a lead time of 3 to 6 months, which is crucial for agricultural planning, reservoir management, and authorities' allocation of water resources.*

**Keywords:** Drought Forecast, SPI, NDVI, NDWI, Machine Learning, Wavelet Transform

## 1 Introduction

Among all natural disasters, drought is the most costly environmental catastrophe [1]. As drought is the consequence of precipitation deficiency over an extended period of time, mitigating the detrimental effects of droughts fundamentally lies in the ability to forecast droughts accurately ahead of time to enable the effective planning of water resources.

Due to complex atmospheric processes, accurate drought prediction over a large time span has been one of the biggest challenges in hydrology. In the United States, the U.S. Drought Monitor [2] provides a weekly update of current drought conditions at the national and state levels by publishing an interactive colored-map. However, few uniform drought early warning systems exist globally due to the complexity of the drought process and expensive operational land surface models.

Over the last decade, there has been growing scientific interest in using statistical data-driven methods for drought forecast. Various machine learning algorithms have been investigated, including Autoregressive Integrated Moving Average [3, 4], Artificial Neural Network (ANN) and Support Vector Regression (SVR) [5-8], and Adaptive Neuro-Fuzzy Inference System (ANFIS) [9]. More recently, wavelet analysis has been introduced for analyzing the time series data at different frequency bands, which demonstrates positive impacts in solving a number of problems in water resources when combined with different machine learning algorithms [10-13].

The above studies, however, have been carried out using only precipitation data for drought prediction. Meanwhile, there has been an increasing popularity of using satellite remote sensors for drought condition monitoring, as satellite data are consistently available and nearly continuous in space and time. For example, Y. Gu performed 5-year grassland drought assessment using Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) [14], and L. Wang introduced Normalized Multi-band Drought Index (NMDI) for monitoring soil and vegetation moisture with remotely-sensed satellite data [15].

Considering the strong correlation between satellite-derived vegetation indices and drought conditions, this study focuses on combining multiple drought-related attributes, including precipitation, satellite-derived land cover vegetation indices, and surface discharge, for forecasting long-term drought in the region of Central Valley, California. The research first investigates how to characterize drought behavior from different monitoring sources, and how to properly combine those attributes for statistical analysis. The wavelet transform (WT) is incorporated with both the ANN and SVR algorithms for drought forecasting. The results are presented for a forecast lead time of up to six months. The performance is analyzed and future improvement is discussed in the end.

## 2 Study area and methodology

This study aims to forecast drought in the region of Central Valley, California. Over the last decade, the California State has suffered severe drought over consecutive months, resulting in significant reduction in groundwater level, lake water capacity, stream-flow, and reservoir storage that were all reflected in the long-term precipitation anomalies. Therefore, this study focuses on the long-term drought forecast associated to the hydrological system. The methodology used in this study is shown in Figure 1.

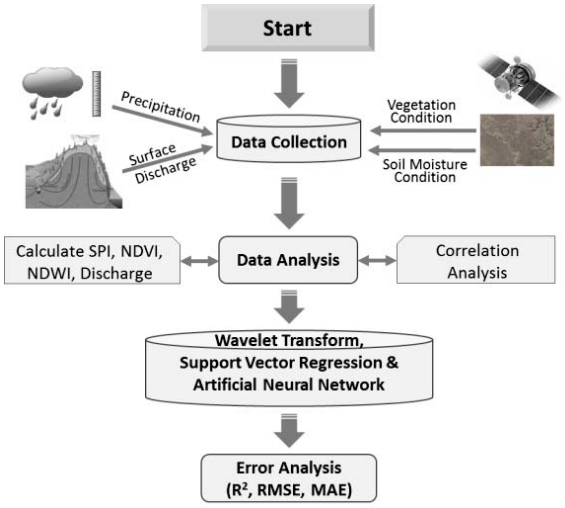


Figure 1: Methodology

The following sections discuss the methods used to characterize drought behavior using the data collected from precipitation, satellite remote sensor and surface discharge. A correlation analysis among those drought attributes is studied for determining the effective inputs to the machine learning algorithms.

## 2.1 Standardized precipitation index calculation

In this study, the Standardized Precipitation Index (SPI) is used for forecasting drought in California, because it allows for temporal flexibility in evaluation of precipitation conditions [16]. Based on the number of months over which the statistical precipitation is calculated, the SPI is defined as SPI3 and SPI6 that represent short-term agricultural drought, and SPI12 and SPI24 that represent long-term hydrological drought. Here SPI3, SPI6, SPI12 and SPI 24 are the SPI for a period of 3, 6, 12, and 24 months respectively. This study focuses on hydrological drought impact so the forecast will be made based on SPI12 and SPI24.

To calculate the SPI, the monthly precipitation from 1948 to 2014 is collected at the two in-situ weather stations close to the cities of Stockton and Sacramento, CA, from the National Oceanic and Atmospheric Administration (NOAA). The concept of SPI12 and SPI24 calculation is described below:

1) Calculate the cumulative precipitation value for SPI12 and SPI24 for each month from 1948 to 2014.

2) Fit the precipitation data for the same month of each year into a Gamma distribution.

3) Convert the Gamma distribution into a standard Gaussian distribution based on an equal probability transformation.

4) The SPI is a z-score and represents an event away from the mean value in Gaussian distribution.

Based on the calculated SPI values at each time scale, the drought can be classified as given in Table I [16].

Table I: SPI value and drought conditions

| SPI Value | Drought Class |
|---|---|
| SPI ≥ 2.0 | Extremely wet |
| 1.5 ≤ SPI < 2.0 | Very wet |
| 1.0 ≤ SPI < 1.5 | Moderate wet |
| -1.0 ≤ SPI < 1.0 | Normal |
| -1.5 ≤ SPI < -1.0 | Moderate drought |
| -2.0 ≤ SPI < -1.5 | Severe drought |
| SPI < -2.0 | Extreme drought |

In this study, the SPI12 and SPI24 are calculated using SPI_SL_6 program developed by the National Drought Mitigation Center, University of Nebraska-Lincoln.

## 2.2 Grassland vegetation indices from satellite remote sensor

Satellite remotely-sensed data are a promising source of drought condition monitoring as it is possible to measure every component of the hydrological cycle at the land surface and the status of natural vegetation and agriculture, at a very high spatial resolution and nearly real time. To calculate the regional satellite-derived indices for surface land vegetation condition assessment, a study area is chosen near Stockton in the Valley, with a grassland cover that allows the influence of drought to be isolated from other human effects. The satellite remotely-sensed data are acquired from the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Aqua and Terra satellites.

Two satellite-derived indices are used to assess vegetation and soil moisture conditions. The Normalized Difference Vegetation Index (NDVI), as defined in Equation (1), is an indication of live green vegetation conditions by detecting the reflection to sunlight at two optical wavelength bands as illustrated in Figure 2. Here $\rho_{645nm}$ and $\rho_{860nm}$ are the reflection detected at 645nm and 860nm respectively. A larger NDVI value means a higher density of green vegetation.

$$NDVI = \frac{\rho_{860nm} - \rho_{645nm}}{\rho_{860nm} + \rho_{645nm}} \quad (1)$$

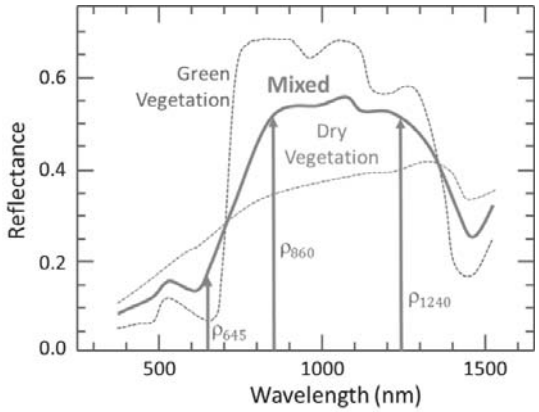


Figure 2: Sunlight reflection vs. wavelength

The Normalized Difference Water Index (NDWI), as defined in Equation (2), uses a similar principle to that for the NDVI to assess soil moisture by detecting the reflection to sunlight at other two optical wavelength bands:

$$NDWI = \frac{\rho_{860nm} - \rho_{1240nm}}{\rho_{860nm} + \rho_{1240nm}} \quad (2)$$

where $\rho_{860nm}$ and $\rho_{1240nm}$ denote the reflection detected at 860nm and 1240nm respectively. The satellite data acquired in this study are 8-day composite of 500-meter surface reflectance data from 2000 to 2014. Data from different optical bands are extracted for multiple pixels. The NDVI and NDWI for each month are calculated according to Equations (1) and (2).

## 2.3 NDVI12/24 and NDWI12/24 calculation

To utilize satellite-derived indices for drought forecasting, the monthly NDVI and NDWI are converted in the same time scale as SPI12 and SPI24 so that they effectively represent the drought conditions to be predicted. First, the moving average is applied to both NDVI and NDWI at each month over its previous 12-month or 24-month separately. Due to the fact that the satellite data are only available for 14 years, which is not long enough to represent an effective statistical distribution, a simple deviation from the mean value of the overall 14 years is calculated for each month. The newly calculated time series are referred to as NDVI12, NDVI24, NDWI12 and NDWI24 respectively. The time series data for SPI12, NDVI12 and NDWI12 are depicted in Figure 3, and SPI24, NDVI24, NDWI24 are plotted in Figure 4.

The surface discharge or stream-flow data from 2000 to 2014 are obtained from the United States Geological Survey (USGS) gage station located in a natural stream near Stockton. A similar method as described above is applied to the discharge data for calculating Discharge12 and Discharge24, which are newly introduced indices based on surface discharges of 12 months and 24 months.

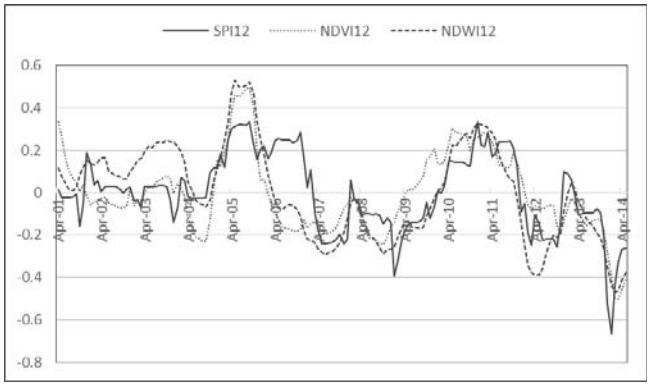


Figure 3: SPI12, NDVI12 and NDWI12 time series plot

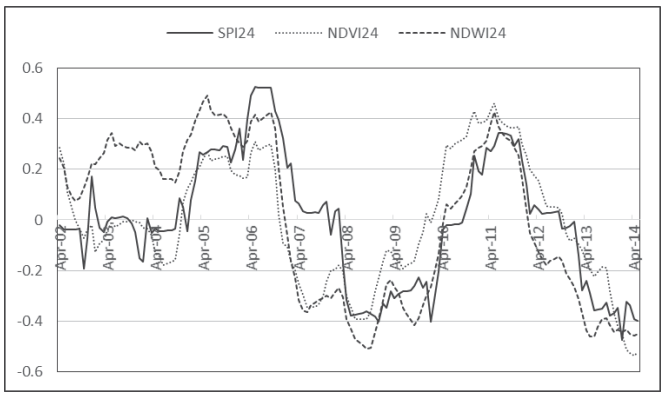


Figure 4: SPI24, NDVI24 and NDWI24 time series plot

## 2.4 Correlation analysis

Before applying NDVI, NDWI and Discharge for augmenting SPI forecast, a correlation analysis is made to ensure that the data taken for the study area are well correlated with the SPI observation data such that they can be helpful for drought prediction. The correlation results are shown in Table II and Table III respectively.

Table II: Correlation coefficient (R) between SPI12 and NDVI12, NDMI12, and Discharge12

| R | NDVI12 | NDWI12 | Discharge12 |
|---|---|---|---|
| SPI12 | 0.62 | 0.72 | 0.57 |

Table III: Correlation coefficient (R) between SPI24 and NDVI24, NDMI24, and Discharge24

| R | NDVI24 | NDWI24 | Discharge24 |
|---|---|---|---|
| SPI24 | 0.73 | 0.78 | 0.69 |

Table II and Table III clearly indicate that NDVI, NDWI and Discharge all have a reasonably good correlation with the SPI, therefore are useful for augmenting the SPI forecast.

# 3 Wavelet-ANN and wavelet-SVR

This section discusses the model development using the ANN and SVR, as well as the method used to apply wavelet transform for data pre-processing. The complete data sets from 2000 to 2014 are divided into two sections: the data from 2000 to 2010 are used for training in each machine learning algorithm; the data from 2011 to 2014 are used for validating the performance of the models.

## 3.1 Artificial neural network

The ANN is a machine learning method which was inspired by how neurons communicate in the human brain. The ANN architecture used in the present study is a feed-forward hierarchical structure that consists of an input layer with multiple input elements, a hidden layer with multiple neurons, and an output layer called the target layer. The ANN used in this study can be represented by [8]:

$$\hat{Y}(t) = f_o\left[\sum_{j=1}^{M} W_j \cdot f_n\left(\sum_{i=1}^{N} W_{ji} \cdot X_i(t) + W_{jO}\right) + W_O\right] \quad (3)$$

where M is the number of neurons in the hidden layer, N is the number of input attributes, i is the input element, j is the hidden neuron, and t is the function of time. $X_i$, $W_{ji}$, $f_n$, $f_o$ and $\hat{Y}$ represent the input viable, the weight operators, the activation function of hidden neuron, the activation function of output neuron, and the forecast output respectively. A recursive, multi-step neural network approach is chosen such that the forecast performance for each leading month can be optimized, as illustrated in Figure 5.

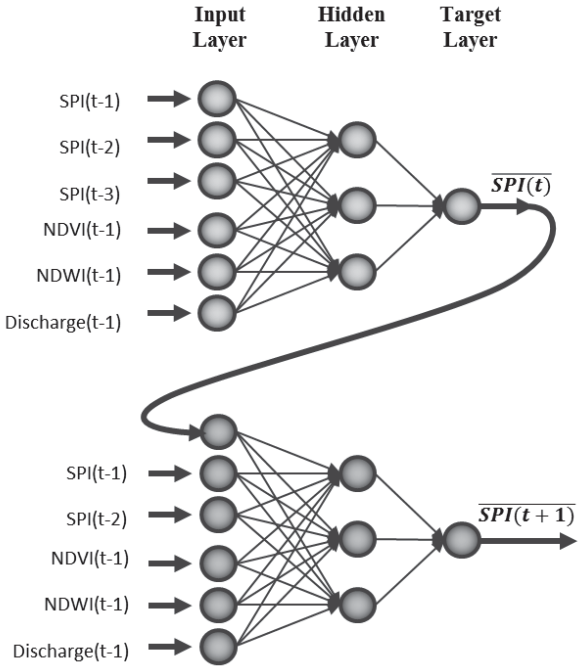


Figure 5: ANN with recursive architecture

In Figure 5, the ANN is trained using Levenberg–Marquardt (LM) algorithm with back propagation [7]. For each leading month forecast, a program is developed using Matlab to optimize the following parameters:

- Number of input combinations
- Number of neurons in the hidden layer
- Activation function in the hidden layer

The optimized ANN model is chosen, which gives the highest determination coefficient ($R^2$) for the validation data set.

## 3.2 Support vector regression

The SVR is a novel machine learning algorithm characterized by the usage of the kernel function, ε-insensitive loss function, and capacity control obtained by a trade-off between the margin maximization and the smoothness of the function f(x) [13]. To solve a non-linear regression problem, the input data space is firstly mapped onto an m-dimensional kernel-induced feature space where linear regression can be applied:

$$f(x) = \sum_{j=1}^{m} w_j \times \phi_j(x) + b \quad (4)$$

where $w_j$ is the weight factor, b is the bias term, $\phi_j$ denotes a set of non-linear transformation functions in the feature space. The goal of the SVR algorithm is to estimate the regression function f(x) that minimizes the following:

$$C\frac{1}{n}\sum_{i=1}^{n} L_\varepsilon(f(x_i), y_i) + \frac{1}{2}\|w\|^2 \quad (5)$$

$$L_\varepsilon(f(x), y) = \begin{cases} 0, & |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon, & |f(x) - y| \geq \varepsilon \end{cases} \quad (6)$$

Equations 5 and 6 describe the trade-off between the empirical risk and the flatness. The term $L_\varepsilon(f(x), y)$ is called ε-insensitive loss function, where y denotes the observed data with a total set number of n, and ε controls the width of the ε-insensitive zone, which can affect the number of support vectors used to construct the regression function. $C\frac{1}{n}\sum_{i=1}^{n} L_\varepsilon(f(x), y)$ is called an empirical error, which measures the deviation of the training samples outside of the ε-insensitive zone. C is the control capacity that determines the trade-off between the tolerated empirical error and the flatness of the model. The smoothness of the function is measured by $\frac{1}{2}\|w\|^2$.

In this study the SVR model is developed using ε-SVR function from LibSVM Matlab toolbox (Chang & Lin, 2014). The kernel function uses Radial Basis Function (RBF) characterized by γ [17]. For each leading month forecast, the parameters C, γ, and ε are optimized through a trial-and-error method for getting the best $R^2$ for the validation data set.

## 3.3 Wavelet analysis for data pre-processing

Despite the fact that both the ANN and SVR are powerful in dealing with non-linear hydrologic problems, they both have limitations that input data must be stationary for a reliable operation. By analyzing the SPI12 and SPI24 time series data and their autocorrelation function (ACF), it can be found that SPI12 and SPI24 are not highly stationary, with the ACF showing a slowly decaying sinusoidal behavior over a number of lags. To mitigate this shortcoming, the wavelet analysis is applied for data pre-processing that allows the use of larger time intervals for more precise low-frequency information and shorter time intervals for extracting high-frequency information, thus generating a time-frequency representation of the time series signal. Using the wavelet analysis algorithm, the original time series data can be hierarchically decomposed into N-level sub-series at different frequency bands for noise reduction or peak detection. This technique is especially useful for analyzing time domain waveforms where sharp spikes need to be localized, which appear to be the case for data plotted in Figure 3 and Figure 4.

In this study, the discrete wavelet transform (DWT) is used to decompose each of the input attributes to the ANN or

SVR models into a number of sub-series at different resolution levels. The Daubechies mother wavelet is used, which provides a family of wavelets called dbN, where N is the order of wavelets. The data decomposition is done by passing the input data through a series of high-pass and low-pass filters, to obtain the detailed series (D) and approximation series (A), as illustrated in Figure 6.

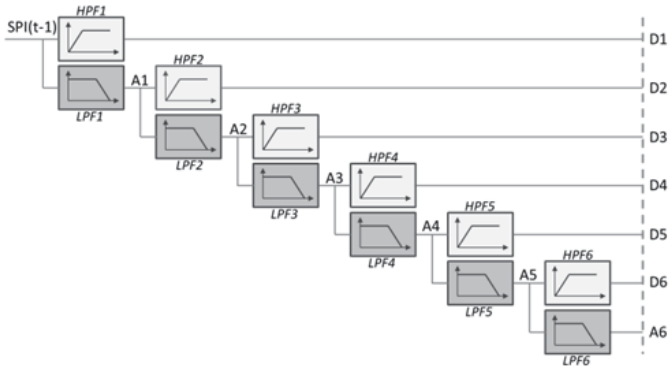


Figure 6: 6-level WT decomposition for SPI(t-1)

As an example, Figure 7 shows the SPI12(t-1) decomposed waveform for D1-D6 and A6, using Daubechies db1 as the master wavelet.

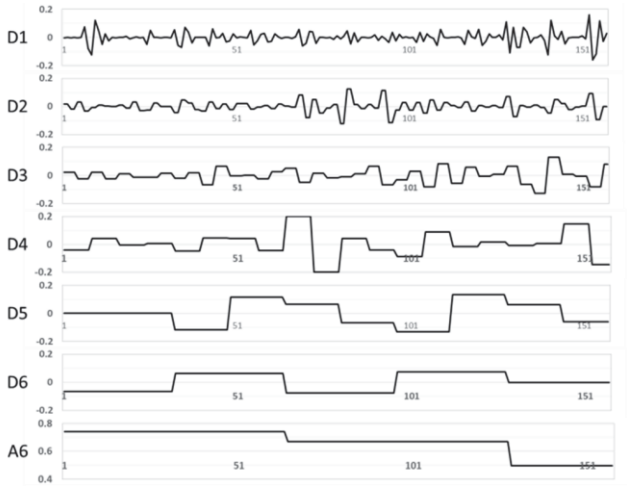


Figure 7: SPI12 decomposed waveforms using DWT

After data decomposition, a portion of the signal associated with certain frequency bands will be eliminated if there is a poor correlation between the decomposed signal and the observation data. Only the decomposed signals that have significant correlation with the observation signal will be used in the forecast model. Figure 8 shows the methodology of using the WT-ANN (WT based ANN) and the WT-VSR (WT based VSR) for forecasting SPI12 and SPI24.

After a correlation analysis, it is found that the newly formed time series data using the DWT algorithm appear to have a better correlation with the observation data than the original time series. Therefore using the DWT for data pre-processing should be helpful for improving SPI12 and SPI24 prediction accuracy. For each leading time prediction using the WT-ANN and WT-SVR, a program is developed using Matlab that optimizes drought forecast performance by using different Daubechies wavelets for data decomposition. The wavelet that gives the highest $R^2$ and the lowest MAE and RMSE will be selected.

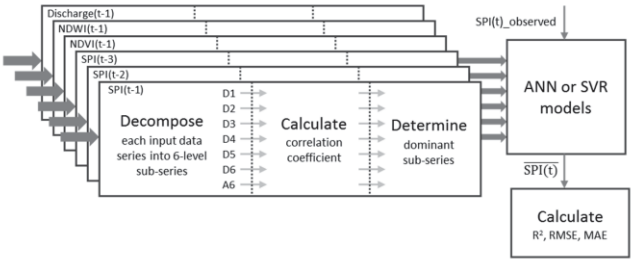


Figure 8: WT-ANN and WT-SVR configuration

# 4 Performance evaluation

To assess the performance of the WT-SVR and WT-ANN models, three statistical performance evaluation criteria are used: $R^2$, RMSE and MAE. The $R^2$ measures the degree of linear correlation between the predicted data and the observed data, the RMSE gives the variant of the total errors, while the MAE provides the absolute error information. The models with the highest $R^2$ and the lowest RMSE and MAE indicate the best performance. The $R^2$, RMSE and MAE are defined in following equations:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (8)$$

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \quad (9)$$

where n is the number of data points, $y_i$ is the observed value, $\hat{y}_i$ is the predicted value, and $\bar{y}_i$ is the mean value of the observation data.

# 5 Results and discussion

Using the algorithms described above, drought forecast based on SPI12 for a lead time of up to 3 months and SPI24 for a lead time of up to 6 months have been investigated using the statistical methods described above. The forecast performances are evaluated on the validation data sets from January 2011 to May 2014, and the results are presented in Table IV and Table V. Figure 9 and Figure 11 illustrate the time series plots for SPI12 and SPI24 using the WT-ANN, including observation, training and validation data sets; Figure 10 and Figure 12 show the corresponding scatter plots for the validation data set, at a 1-month lead time.

Table IV: SPI12 prediction using ANN and SVR

| Leading Time | WT-ANN | | | WT-SVR | | |
|---|---|---|---|---|---|---|
| | $R^2$ | RMSE | MAE | $R^2$ | RMSE | MAE |
| 1-month | 0.91 | 0.29 | 0.21 | 0.91 | 0.29 | 0.21 |
| 2-month | 0.74 | 0.52 | 0.34 | 0.73 | 0.50 | 0.32 |
| 3-month | 0.69 | 0.59 | 0.41 | 0.60 | 0.62 | 0.40 |

Table V: SPI24 prediction using ANN and SVR

| Leading Time | WT-ANN | | | WT-SVR | | |
|---|---|---|---|---|---|---|
| | $R^2$ | RMSE | MAE | $R^2$ | RMSE | MAE |
| 1-month | 0.97 | 0.13 | 0.10 | 0.97 | 0.14 | 0.11 |
| 2-month | 0.95 | 0.19 | 0.12 | 0.95 | 0.19 | 0.15 |
| 3-month | 0.95 | 0.21 | 0.16 | 0.95 | 0.19 | 0.15 |
| 4-month | 0.95 | 0.20 | 0.17 | 0.94 | 0.21 | 0.17 |
| 5-month | 0.93 | 0.24 | 0.18 | 0.93 | 0.27 | 0.21 |
| 6-month | 0.94 | 0.21 | 0.17 | 0.90 | 0.34 | 0.27 |

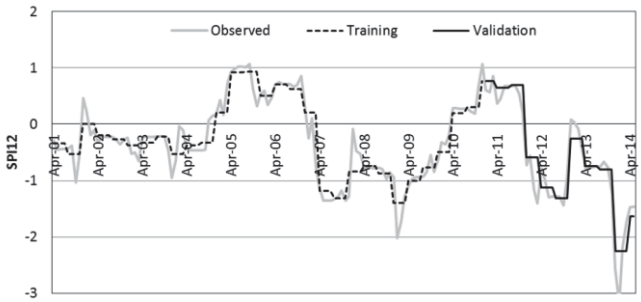


Figure 9: SPI12 time series for observation, training and validation using WT-ANN

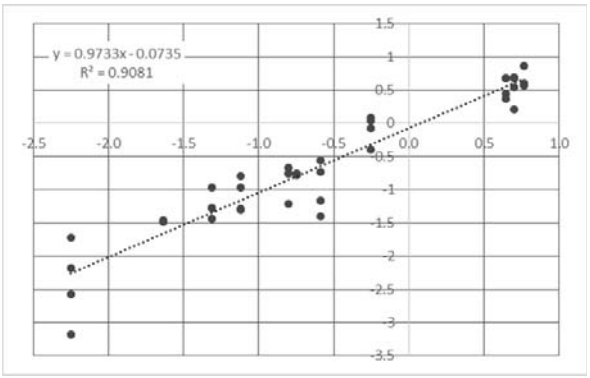


Figure 10: SPI12 scatter plot using WT-ANN

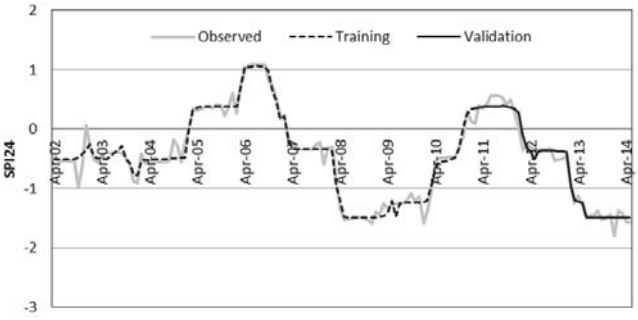


Figure 11: SPI24 time series for observation, training and validation using WT-ANN

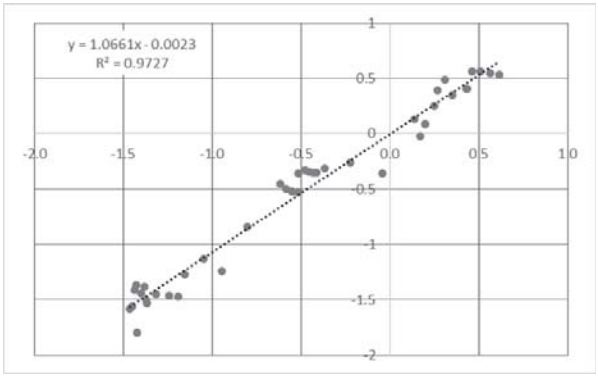


Figure 12: SPI24 scatter plot using WT-ANN

To demonstrate the impact of using satellite data on drought forecast, Figure 13 and Figure 14 illustrate the results of comparison between analyses with and without satellite data, when the same WT-ANN model is used.

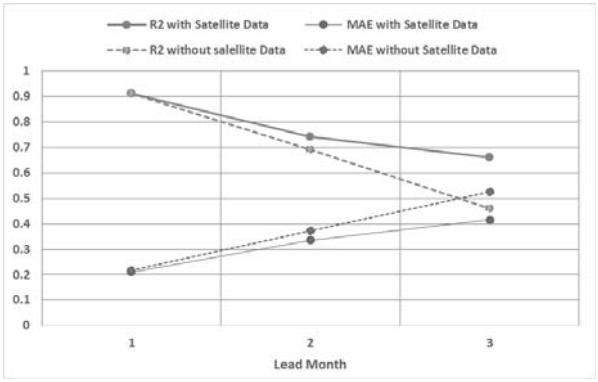


Figure 13: SPI12 forecast analyses with and without satellite data

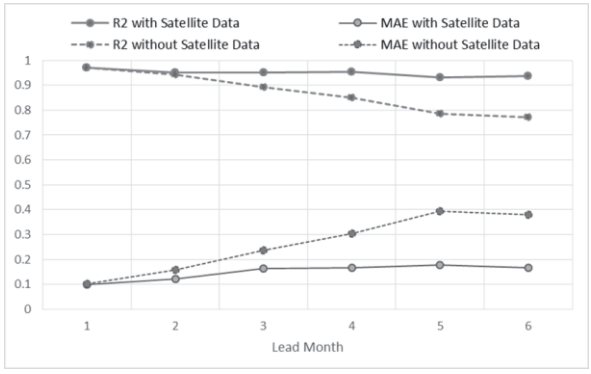


Figure 14: SPI24 forecast analyses with and without satellite data

The aforementioned results indicate the following:

1) Using both the WT-ANN and WT-SVR models, the inclusion of satellite data provides a better accuracy than the results obtained with precipitation data only. The effect of hybrid precipitation and satellite data becomes more obvious as forecast leading time increases. This can be explained by the partial autocorrelation function analysis which shows a strong correlation between SPI(t) and SPI(t-1), therefore, for

one-month lead time prediction, the SPI(t) will be dominated by SPI(t-1).

2) For the SPI24 forecast, the obtained $R^2$ is greater than 0.9 using both the WT-ANN and WT-SVR methods for a lead time of 1-6 months ahead. The superior forecast accuracy is partially due to the fact that SPI24 represents the average precipitation over a period of 24 months, therefore is less sensitive to the monthly variation in precipitation.

3) For the SPI12 forecast, both the WT-ANN and WT-SVR indicate $R^2$ greater than 0.9 for a one-month lead time, but the results are getting worse as lead time increases. This can be explained by the SPI12 observation, training, and validation plots in Figure 9, which shows sharp spikes in multiple time slots that are not well predicted even for one-month lead time. This error will be accumulated as lead time increases. In fact, using Daubechies wavelet for data pre-processing has helped greatly in reducing the sensitivity to big changes in monthly precipitation within the SPI12. Further improvement of SPI12 forecast can be investigated by using other types of wavelet transform or parameter optimization of machine learning algorithms.

# 6 Conclusions

This study aims to investigate the ability of machine learning methods for long-term hydrological drought forecast in the region of Central Valley, CA, based on SPI12 and SPI24. A solution is proposed, for the first time, using hybrid precipitation and satellite remotely-sensed data for drought forecast. The results indicate that integrating precipitation and remotely-sensed data is a promising solution for an effective drought forecast. It is also demonstrated that combining wavelet analysis with novel ANN or SVR is a powerful method for improving the drought forecast accuracy, especially when data is not stationary. To enhance the drought prediction capability, especially for SPI12, further work can be carried out including an investigation of other wavelet functions, as well as taking other drought-related attributes into account, such as ground water level, snowpack and soil moisture. A hybrid geographic and statistical method can also be considered for future study [18].

# 7 References

[1] D. A. Wilhite and M. D. Svoboda, "Drought Early Warning Systems in the Context of Drought Preparedness and Mitigation", National Drought Mitigation Center, Lincoln, Nebraska, U.S.A.

[2] "United States Drought Monitor", http://droughtmonitor.unl.edu/

[3] K. Shatanawi, "Characterizing, Monitoring and Forecasting of Drought in Jordan River Rasin", Journal of Water Resource and Protection, pp. 1192-1202, May 2013

[4] J. Adamowski, H. F. Chan, S. O. Prasher, B. Ozga-Zielinski and A. Sliusarieva, "Comparison of Multiple Linear and Nonlinear Regression, Autoregressive Integrated Moving Average, Artificial Neural Network, and Wavelet Artificial Neural Network Methods for Urban Water Demand Forecasting in Montreal, Canada", Water Resources Research, Vol. 48, W01528, 2012

[5] A. K. Mishra, and V. R. Desai, "Drought Forecasting Using Stochastic Models", Stoch Environ Res Risk Assess, pp.326-339, June 2005

[6] A. K. Mishra, V. R. Desai, and V. P. Singh, "Drought Forecasting Using a Hybrid Stochastic and Neural Network Model", Journal of Hydrologic Engineering, pp. 626-638, December 2007

[7] S. Morid, V. Smakhtin, and K. Baghertadeh, "Drought Forecasting using Artificial Neural Networks and Time Series of Drought Indices", International Journal of Climatology, pp. 2103-2111, April 2007

[8] A. Belayneh, and J. Adamowski, "Drought Forecasting using New Machine Learning Methods", Journal of Water and Land Development, pp. 3-12, No. 18, 2013

[9] B. Shirmohammadi, H. Moradi, V. Moosavi, M. T. Semiromi, and A. Zeinali, "Forecasting of meteorological Drought using Wavelet-ANFIS Hybrid Model for Different Time Steps", Springer Science + Business Media Dordrecht, May 2013

[10] J. R. Mohammed, and H. M. Ibrahim, "Hybrid Wavelet Artificial Neural Network Model for Municipal Water Demand Forecasting", ARPN Journal of Engineering and Applied Sciences, pp. 1047-1065, Vol. 7, No. 8, August 2012

[11] A. Belaynch, and J. Adamowski, "Standard Precipitation Index Drought Forecasting Using Neural Networks, Wavelet Neural Networks, and Support Vector Regression", Applied Computational Intelligence and Soft Computing, 2012

[12] A. Belaynch, J. Adamowski, B. Khalil, and B. Ozga-Zielinski, "Long-term SPI Drought Forecasting in the Awash River Basin in Ethiopia Using Wavelet Neural Network and Wavelet Support Vector Regression Models", Journal of Hydrology, pp. 418-429, 2013

[13] Q. Feng, X. Wen and J. Li, "Wavelet Analysis-Support Vector Machine Coupled Models for Monthly Rainfall Forecasting in Arid Regions", Water Resour Manage, # 29, pp1049 – 1065, 2015

[14] Y. Gu, J. Brown, J. P. Verdin, and B. Wardlow, "A Five-year Analysis of MODIS NDVI and NDWI for Grassland Drought Assessment over the Central Great Plains of the United States", Geophysical Research Letters, Vol. 34, L06407, 2007

[15] L. Wang and J. J. Qu, "NMDI: "A Normalized Multi-band Drought Index for Monitoring Soil and Vegetation Moisture with Satellite Remote Sensing", Geophysical Research Letters, Vol. 34, L20405, 2007

[16] "The Standard Precipitation Index (SPI)", http://drought.unl.edu/ranchplan/DroughtBsics/WeatherDrought/MeasuringDrought.aspx

[17] A. J. Smola and B. Scholkopf, "A Tutorial on Support Vector Regression", NeuroCOLT2 Technical Report Series, 1998

[18] H. Yan and F. Edwards, "Effects of Land Use Change on Hydrologic Response at a Watershed Scale, Arkansas", J Hydrol Eng., Vol. 18, #12, pp. 1779–1785, December 2013